

FIG 1

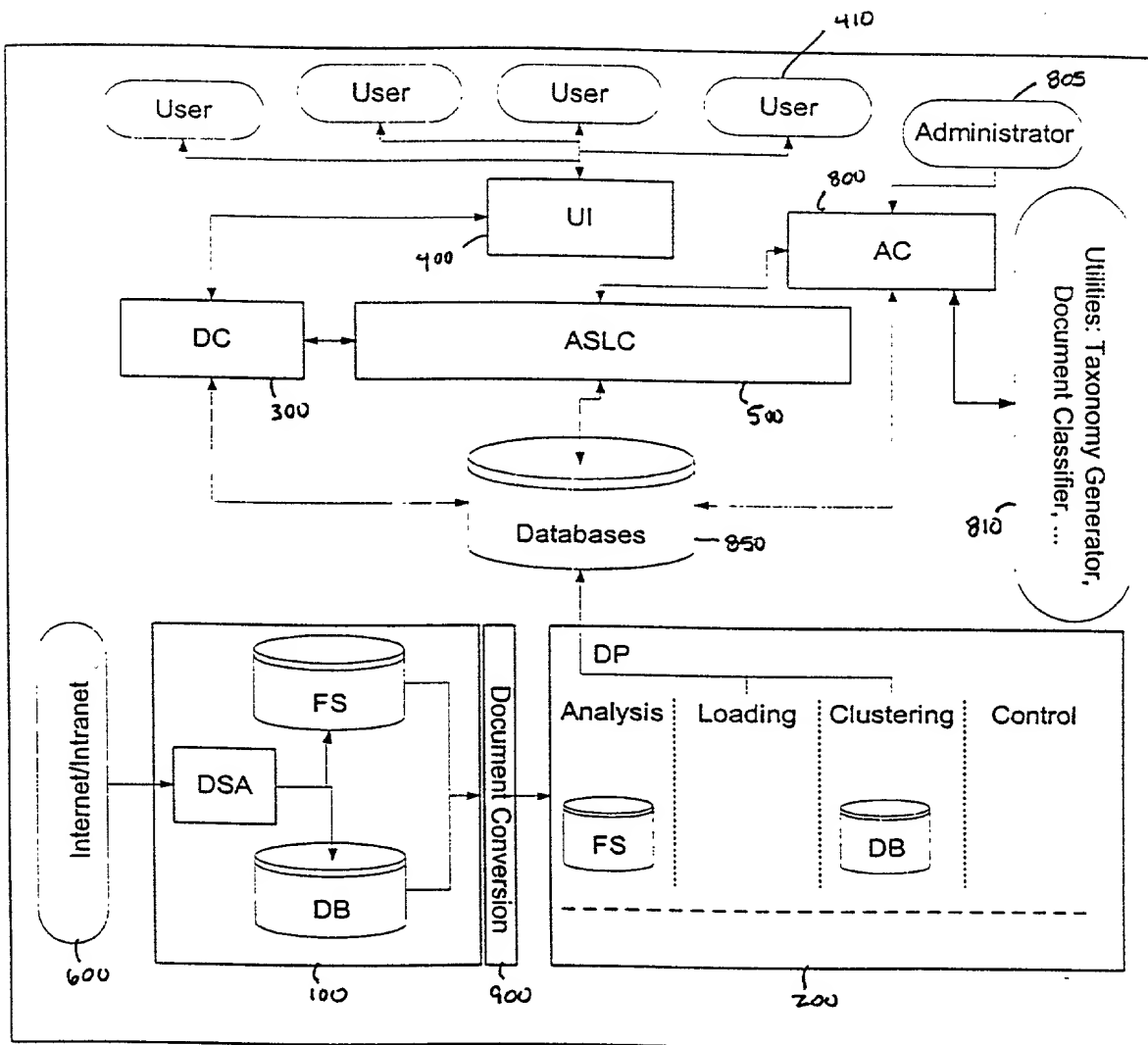


FIG 2

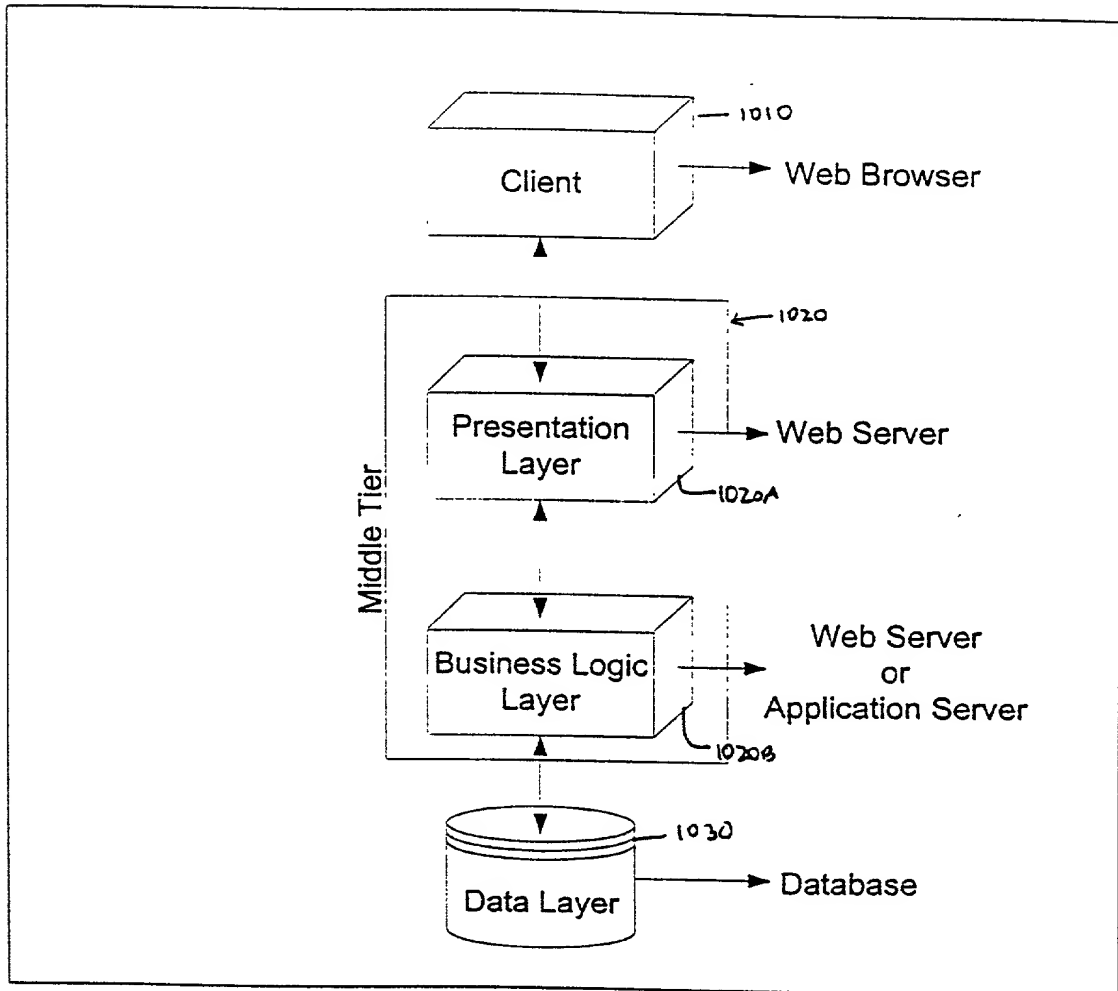


FIG 3

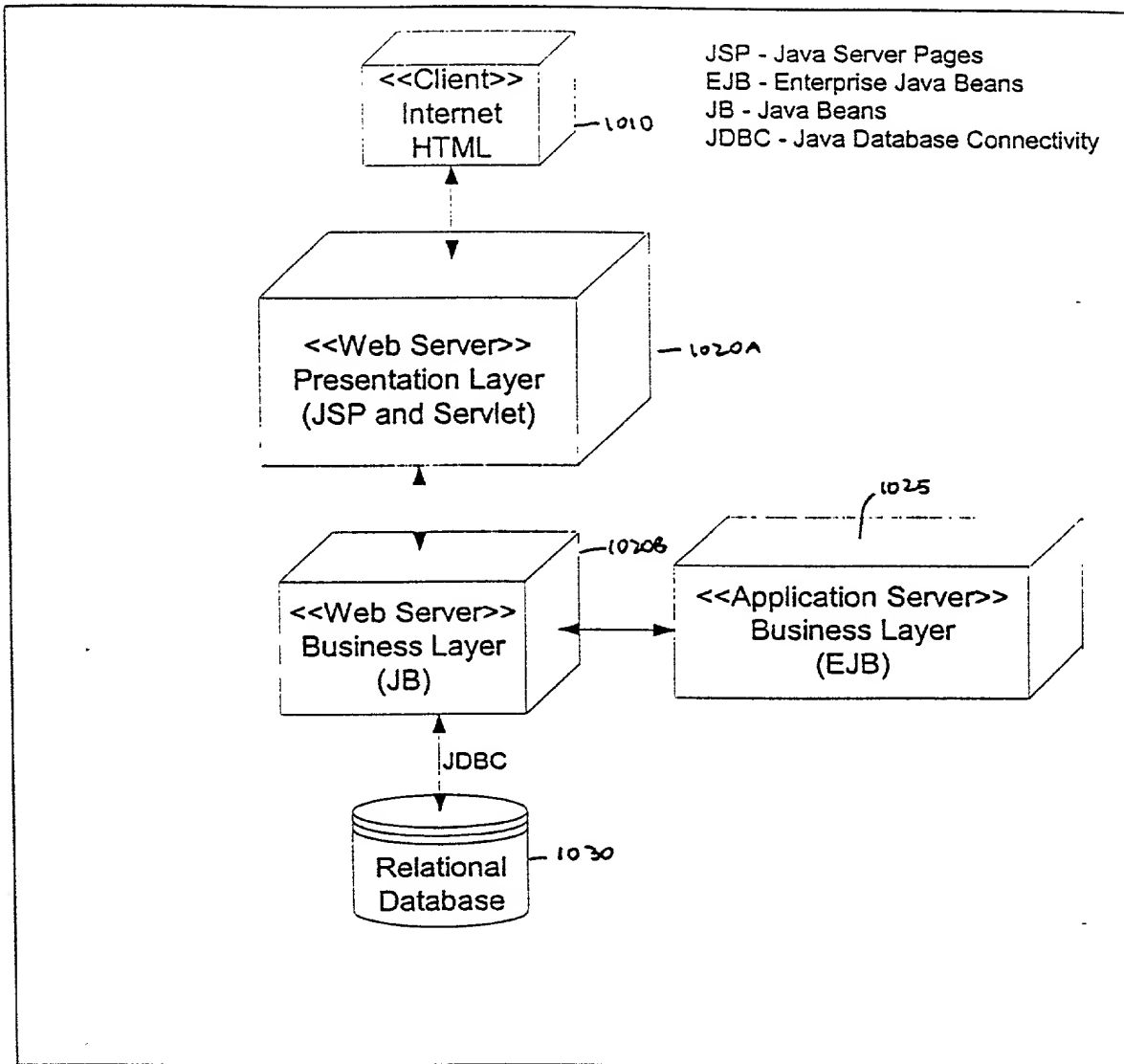


FIG 4

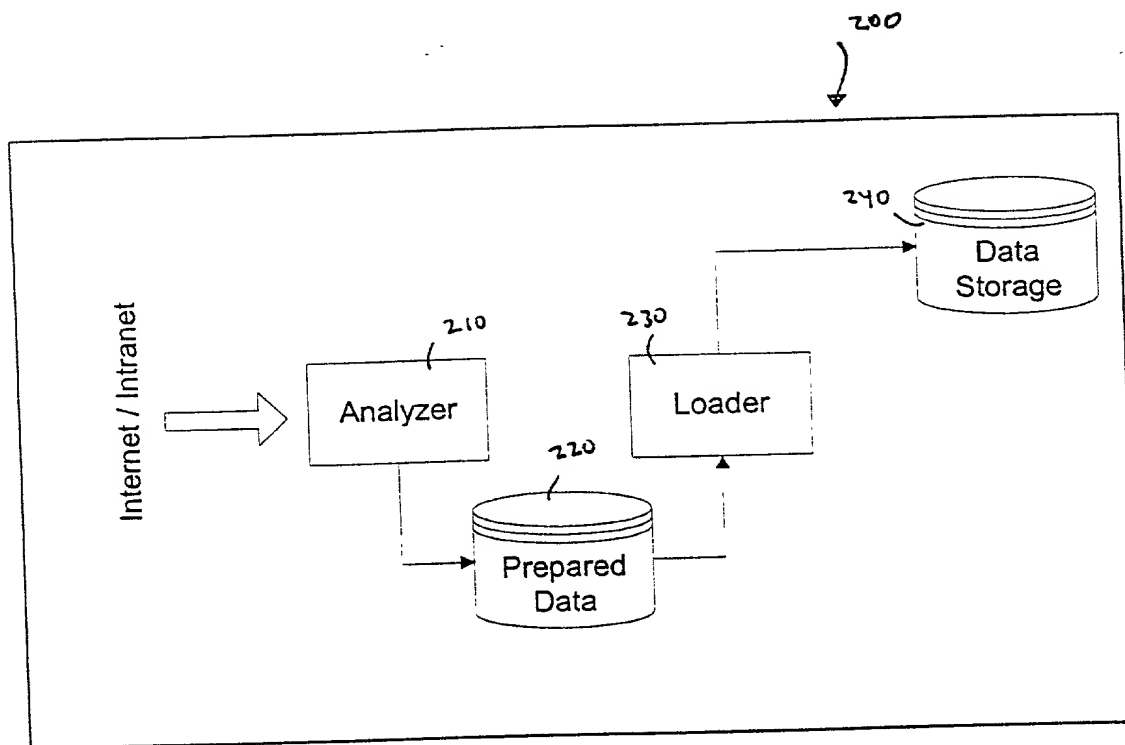


FIG 5

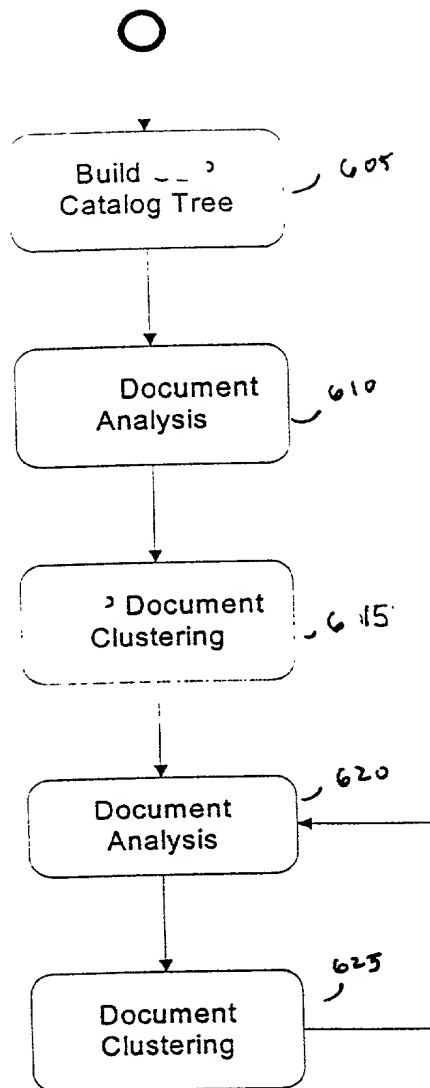


FIG 6

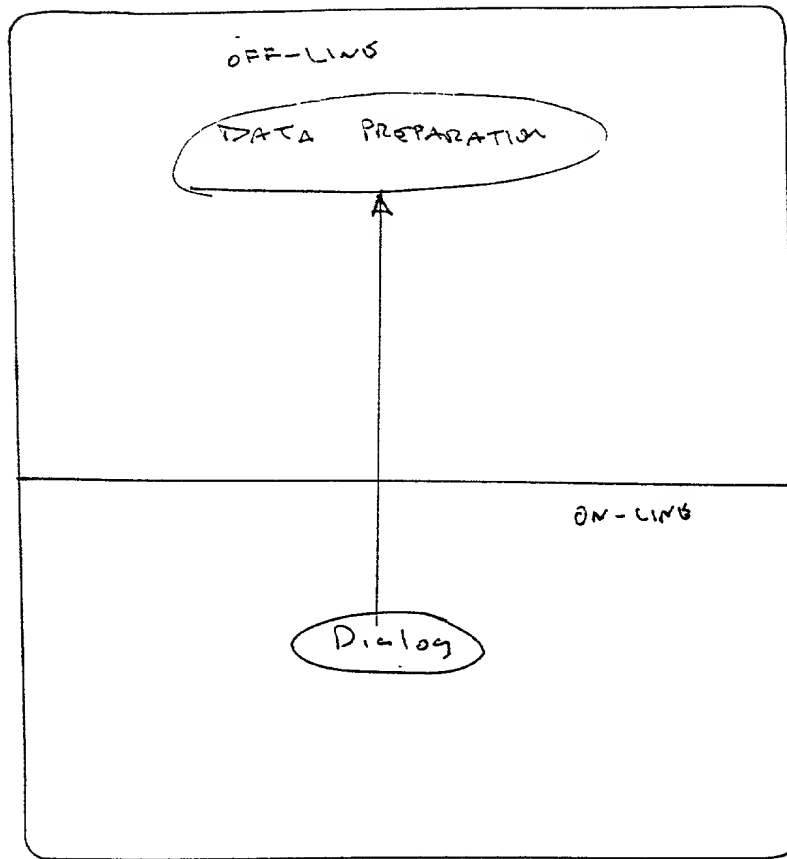


Figure 7

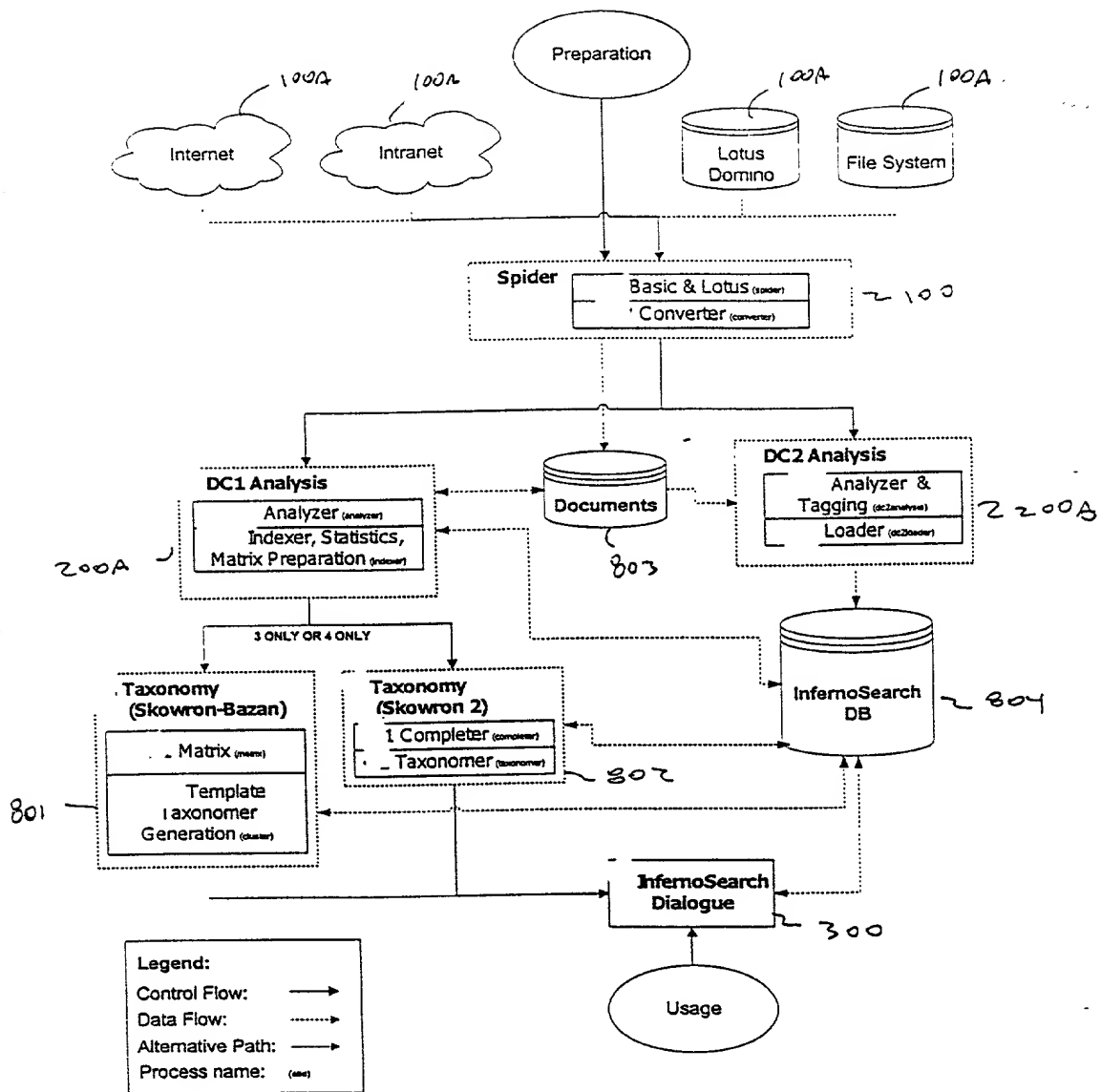


FIGURE 8

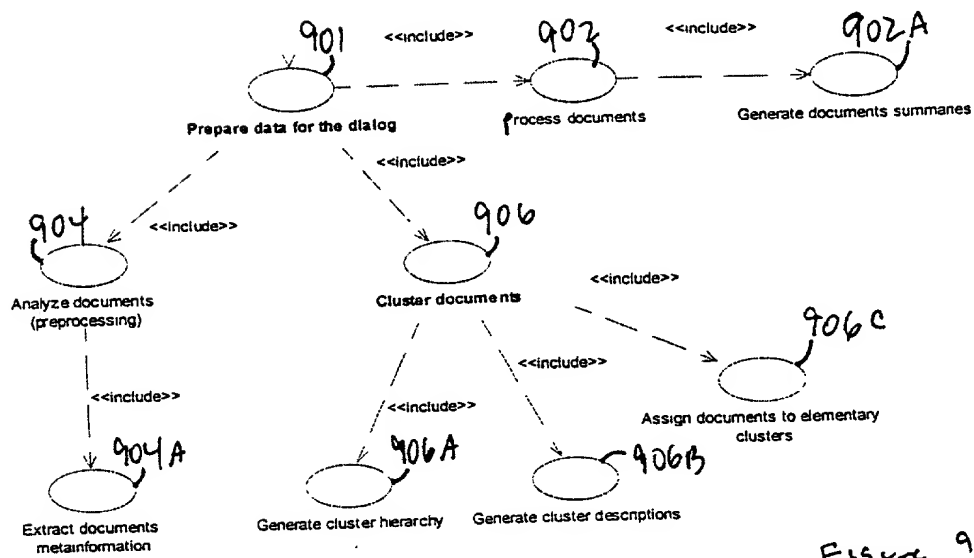
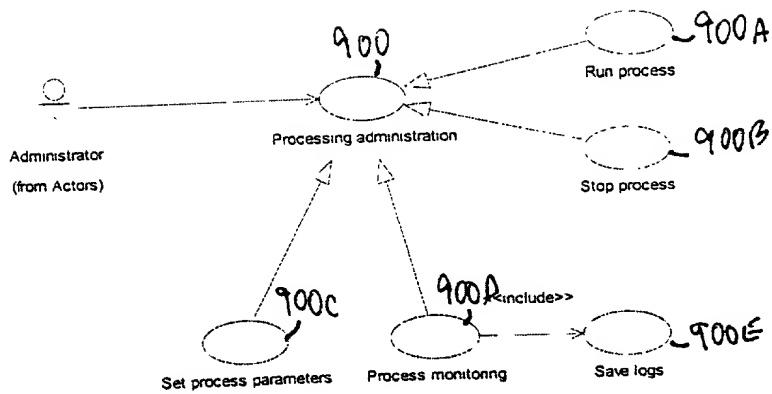


Figure 9

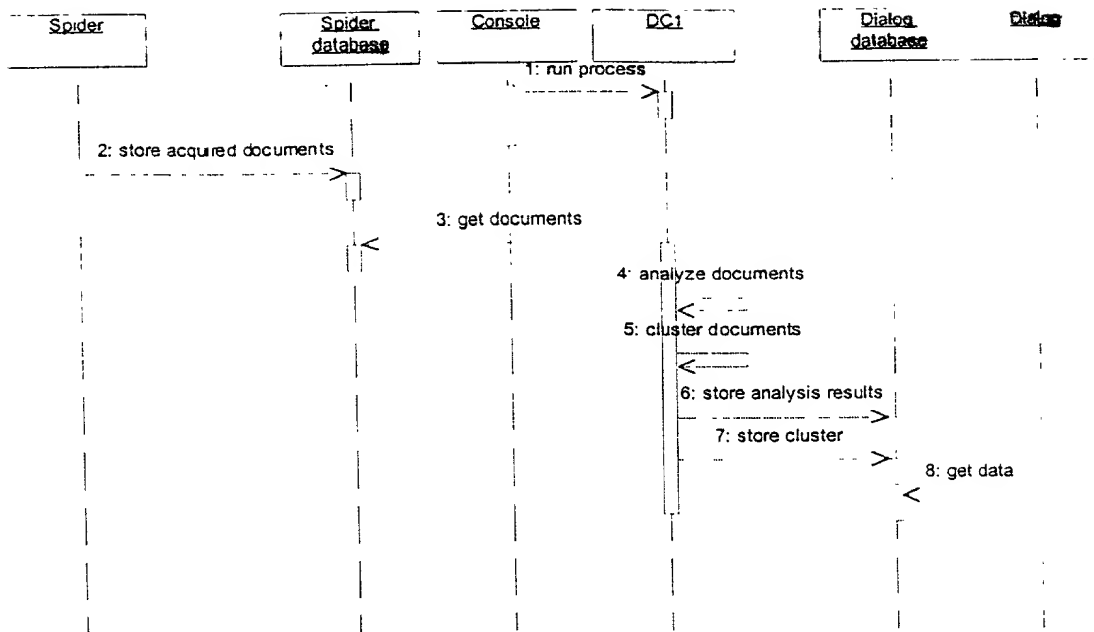


Figure 10

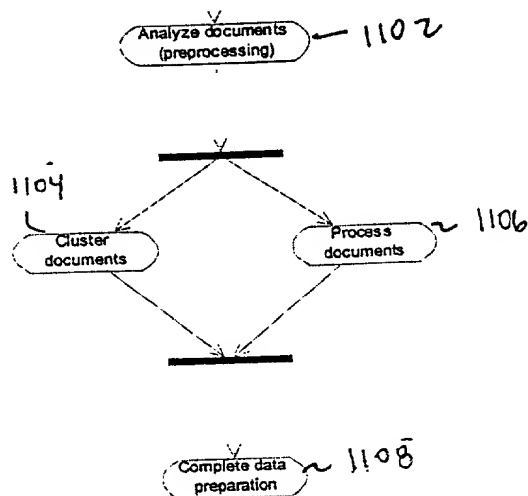


Fig. 11

Figure 12

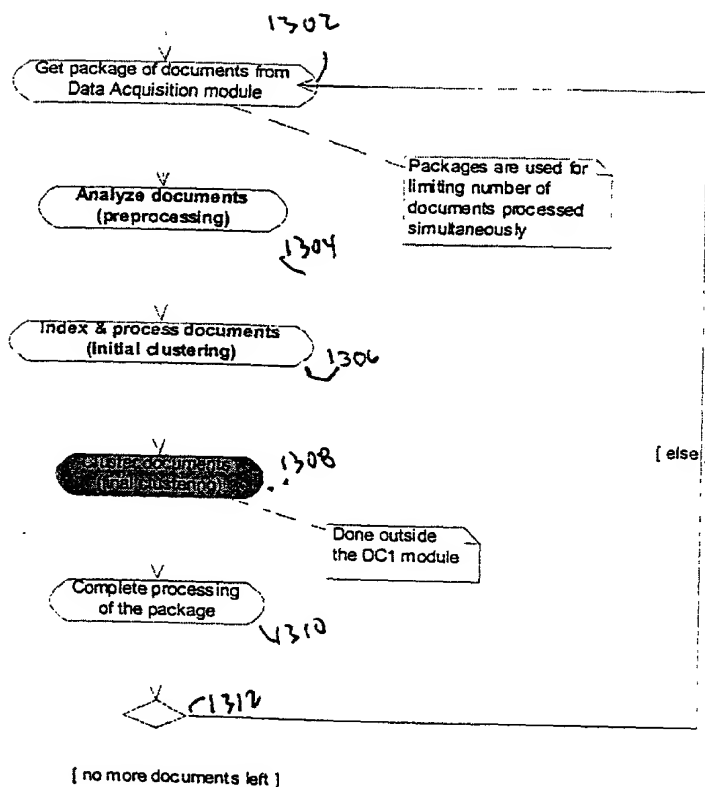
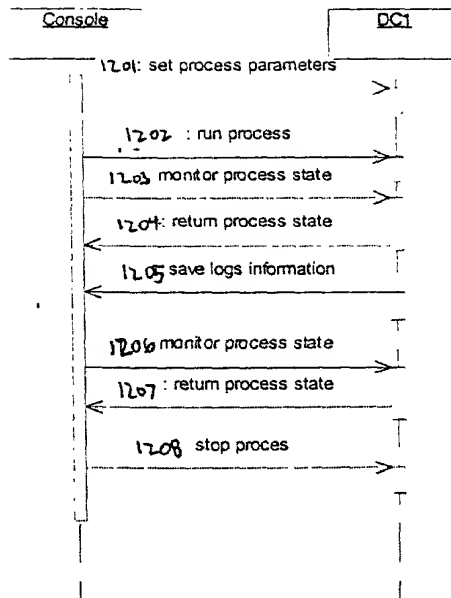


Figure 13

20250323 22:29:50

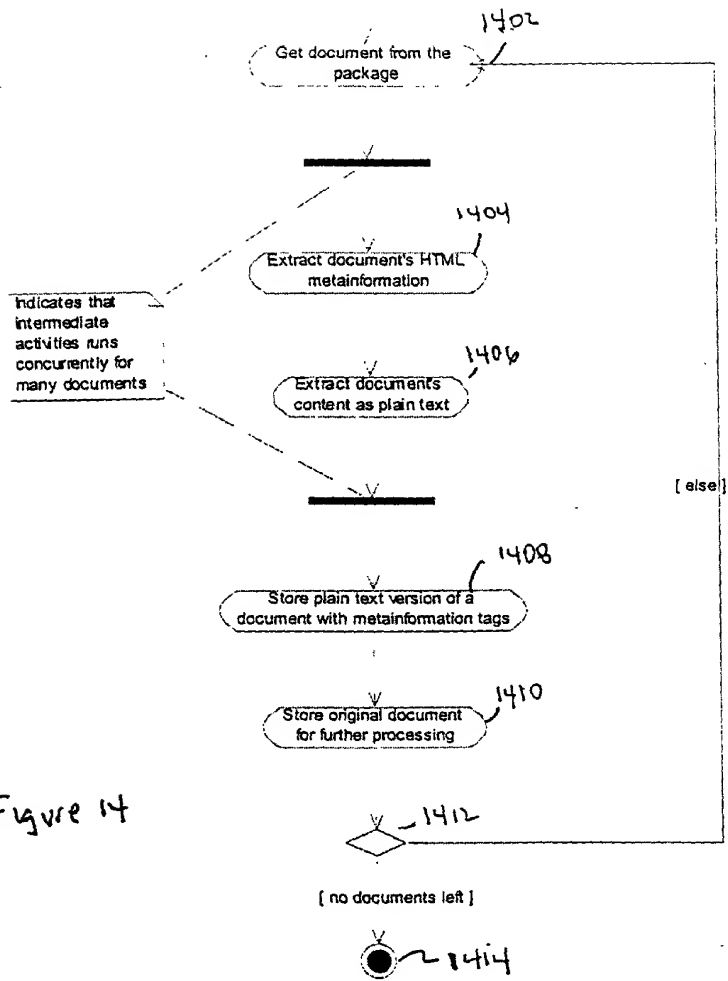


Figure 14

Terms "cluster" and "set of documents containing the same phrase" are equivalent on this diagram

1502
Create local reverted index and dictionary of words/phrases

1504
Generate documents' summaries and prepare statistical data for final clustering

1506
Update global data with local data

1508
Generate clusters hierarchy

1510
Generate clusters descriptions

1512
Find elementary clusters with too many documents

1514
Find documents not covered by any elementary cluster

1516
Complete initial clustering

Figure 15

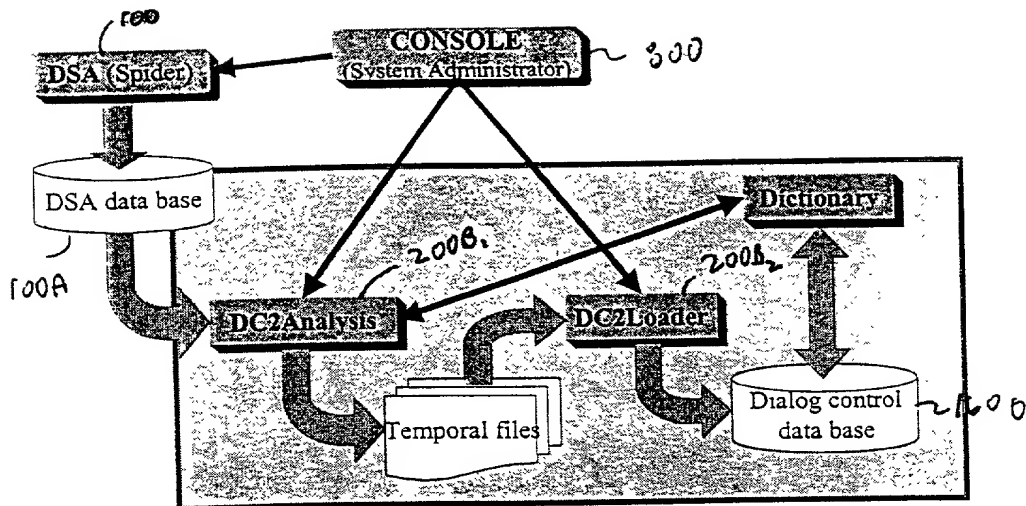


Figure 16

Legends

	Data flow
	Control flow

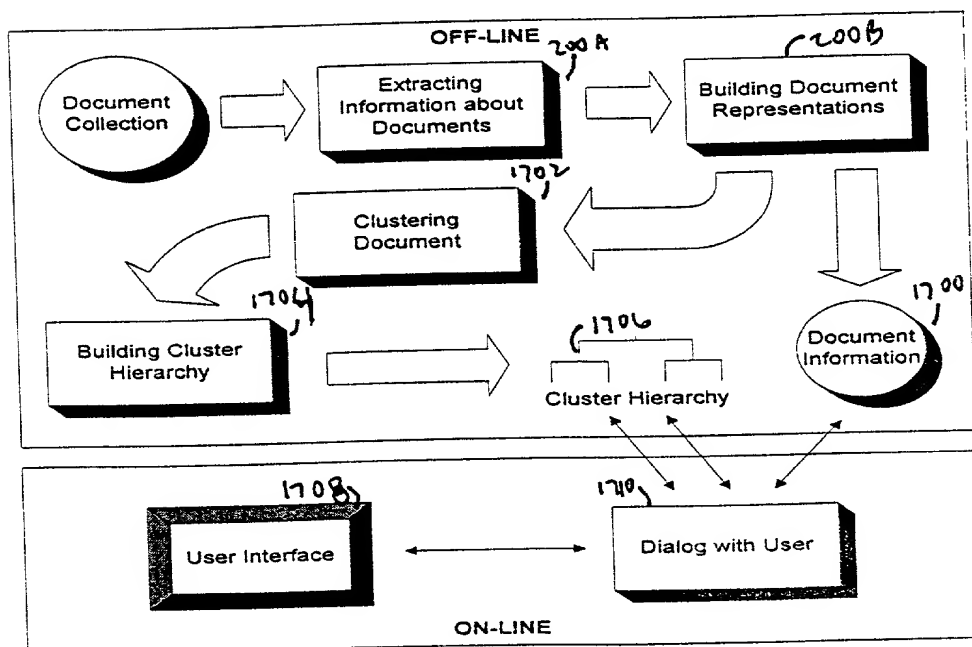


Figure 17

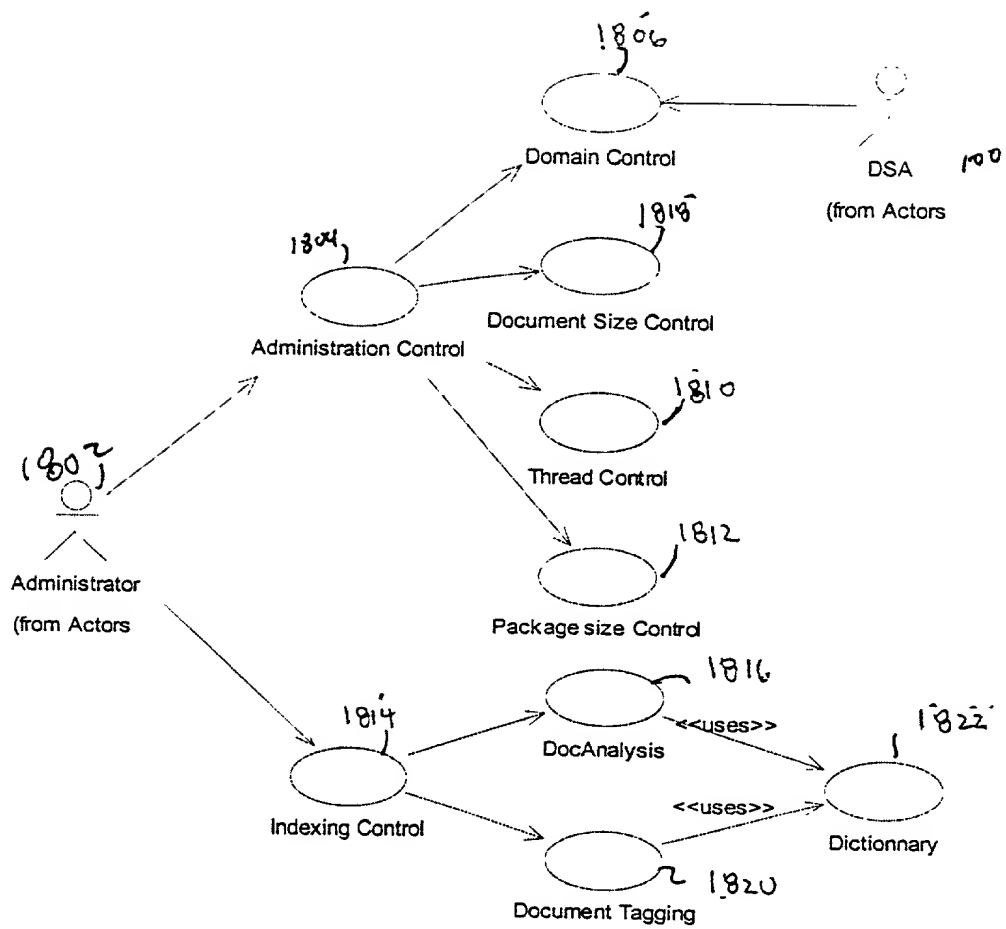


Figure 18

1902 1904 1906 1908 1910 1912 1914 1916 1918

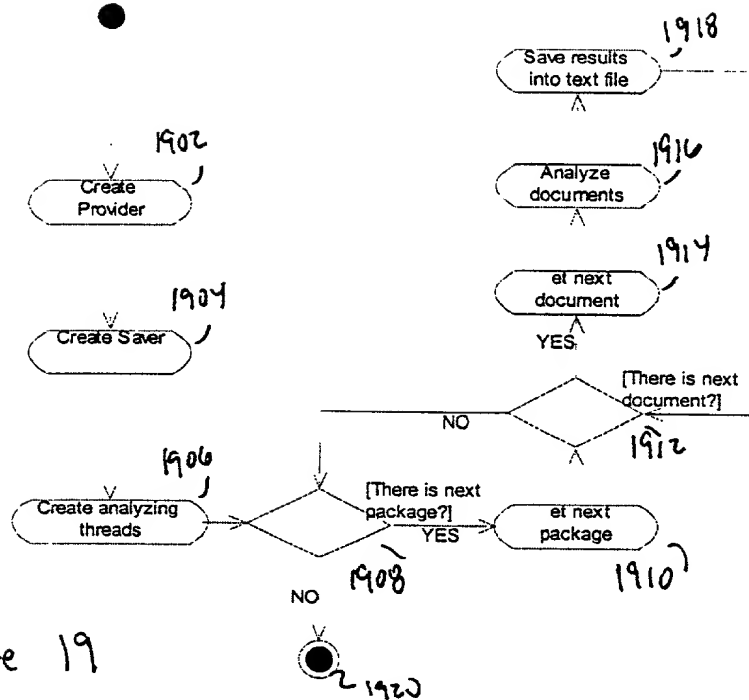


Figure 19

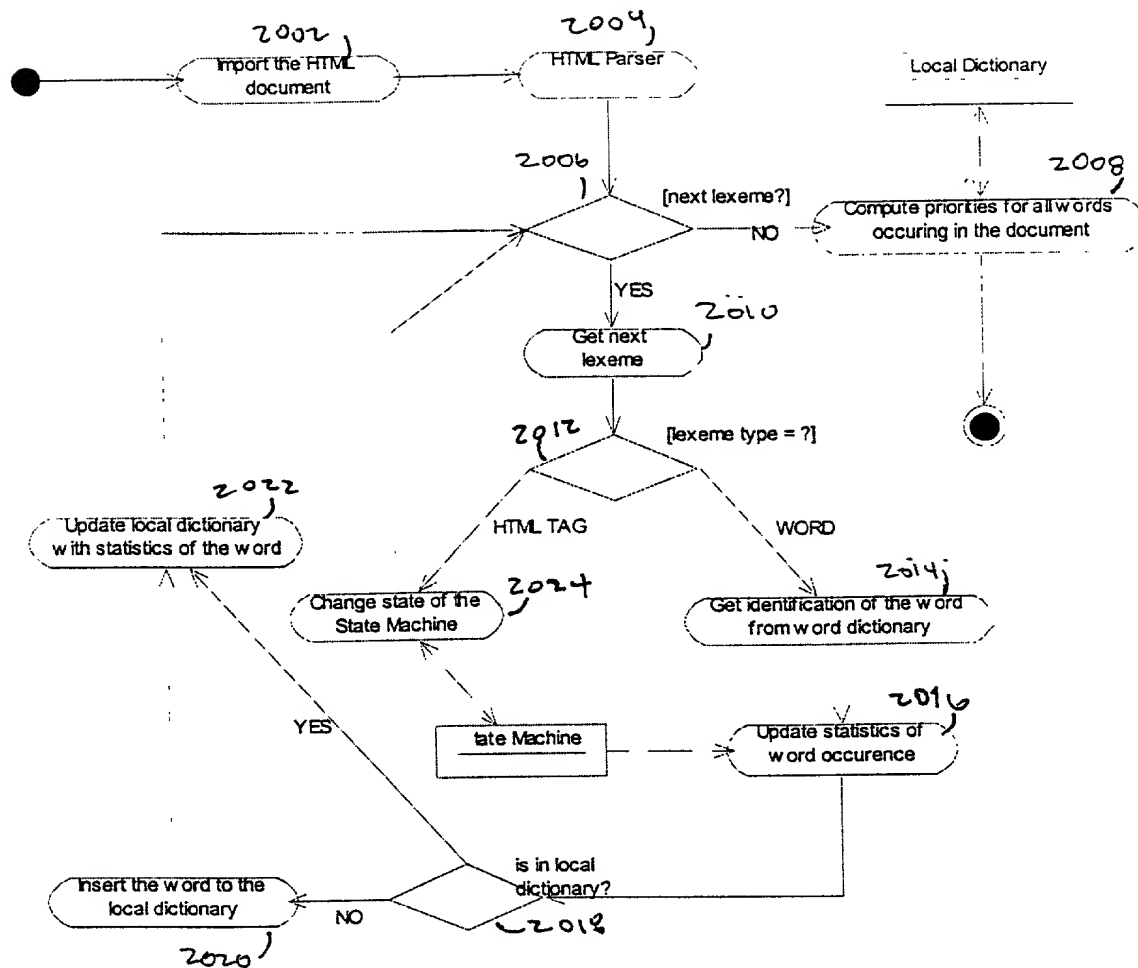


Figure 20

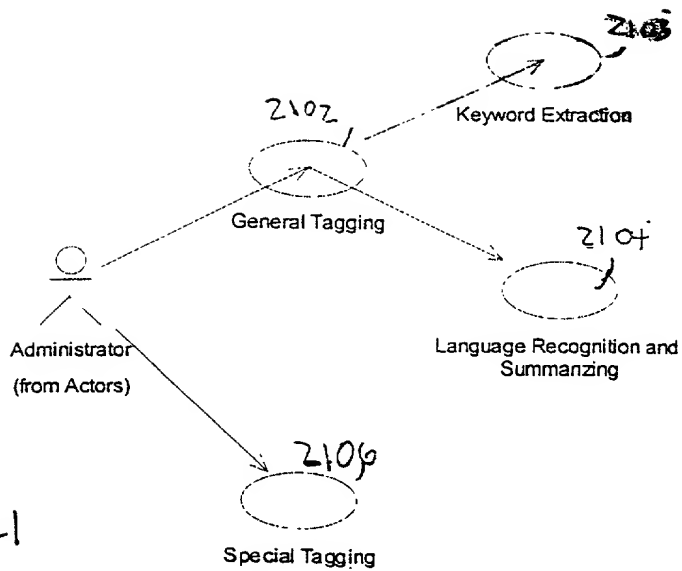


Figure 21

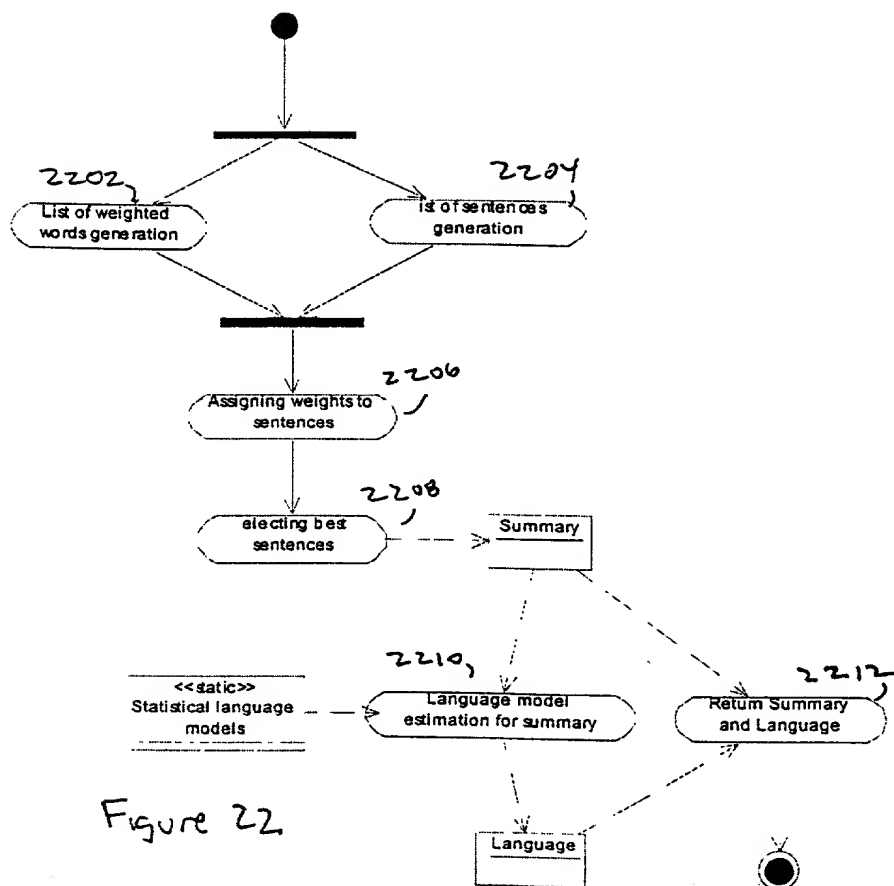


Figure 22

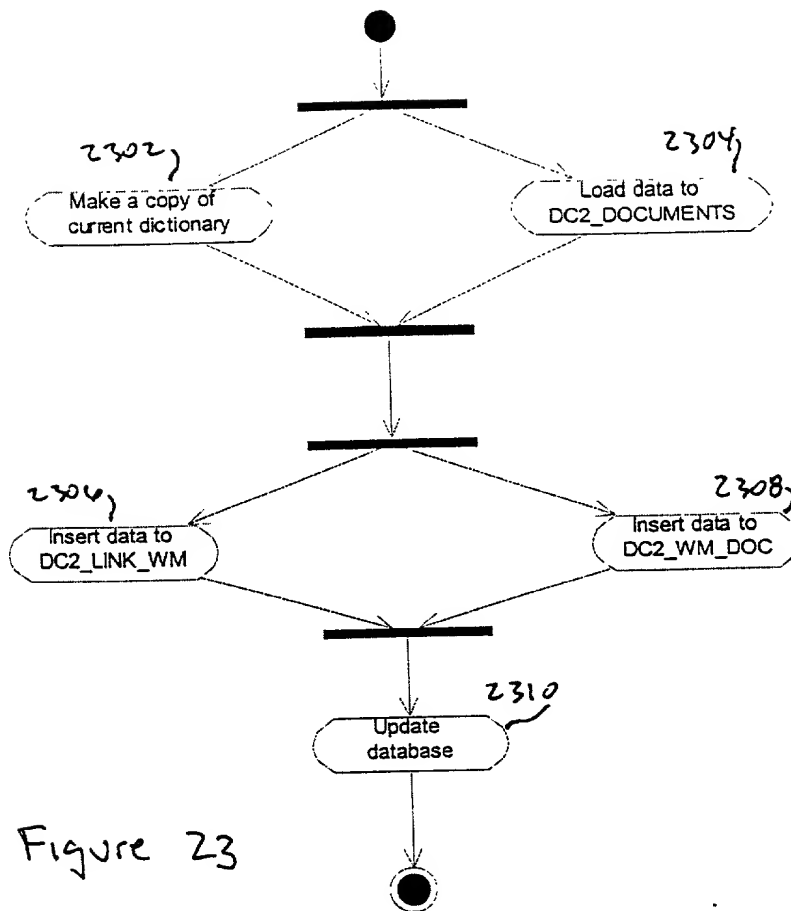


Figure 23

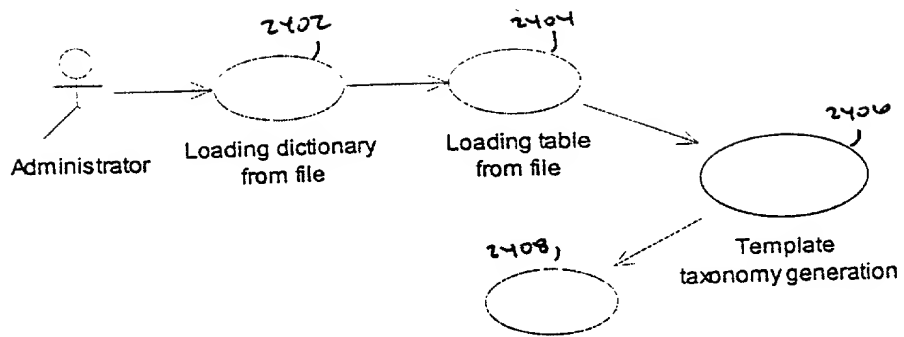


Figure 24

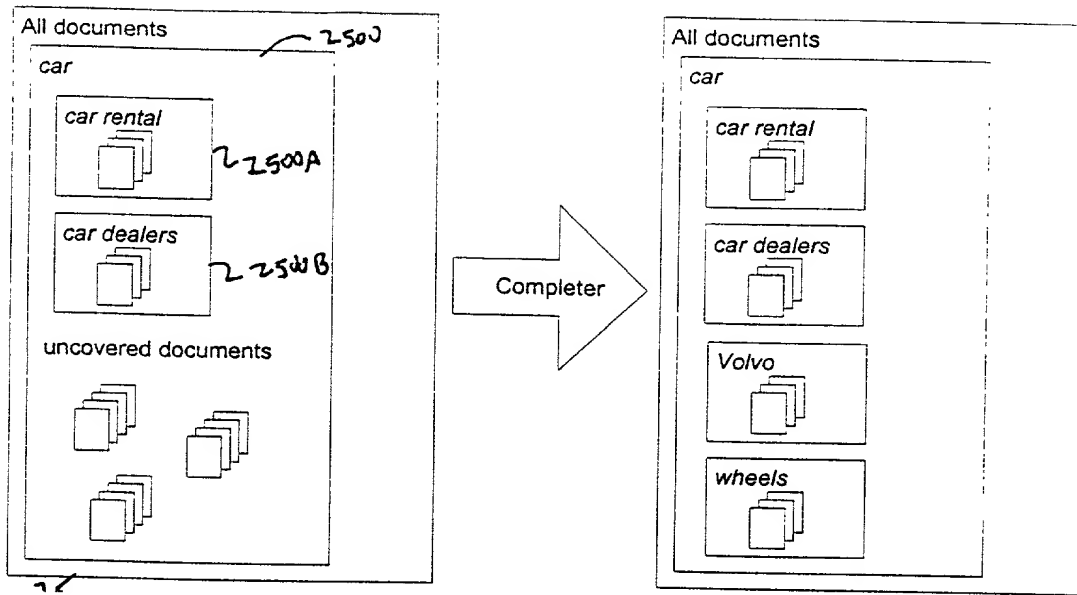


Figure 25

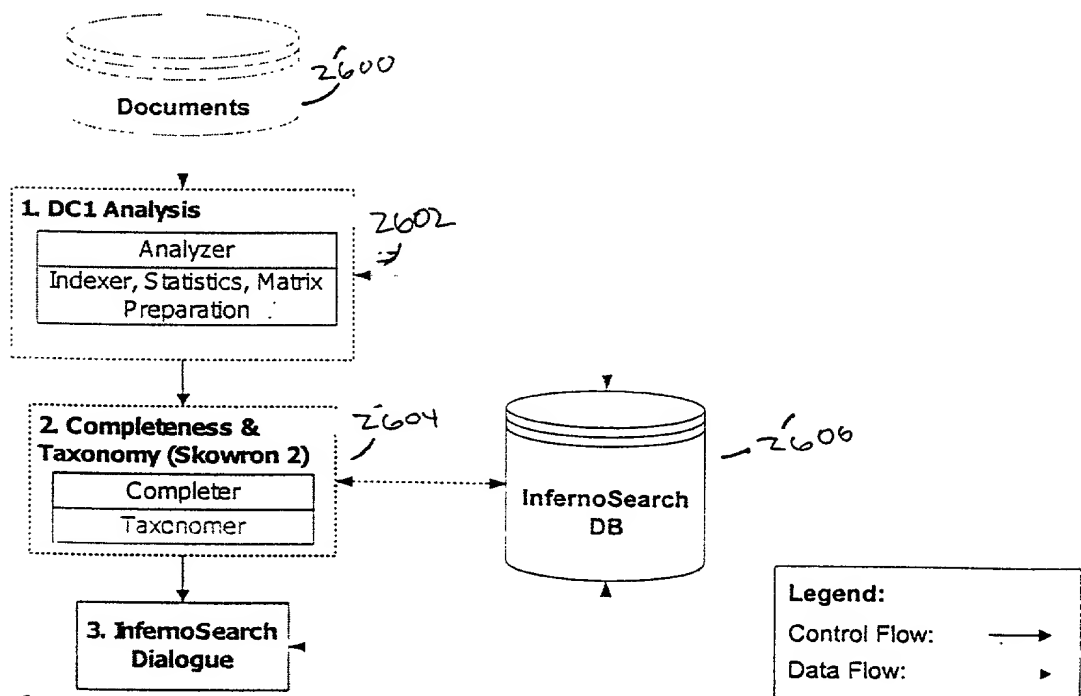
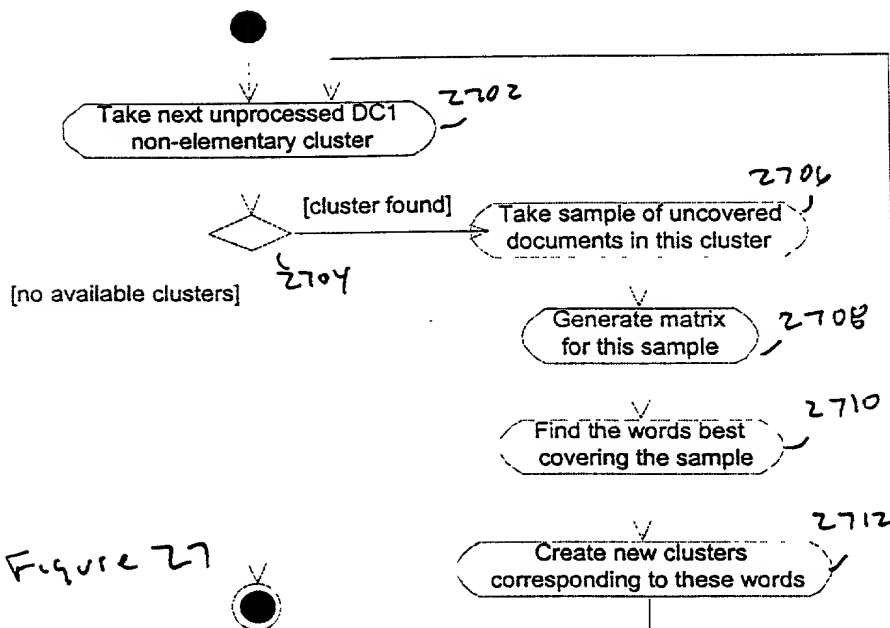
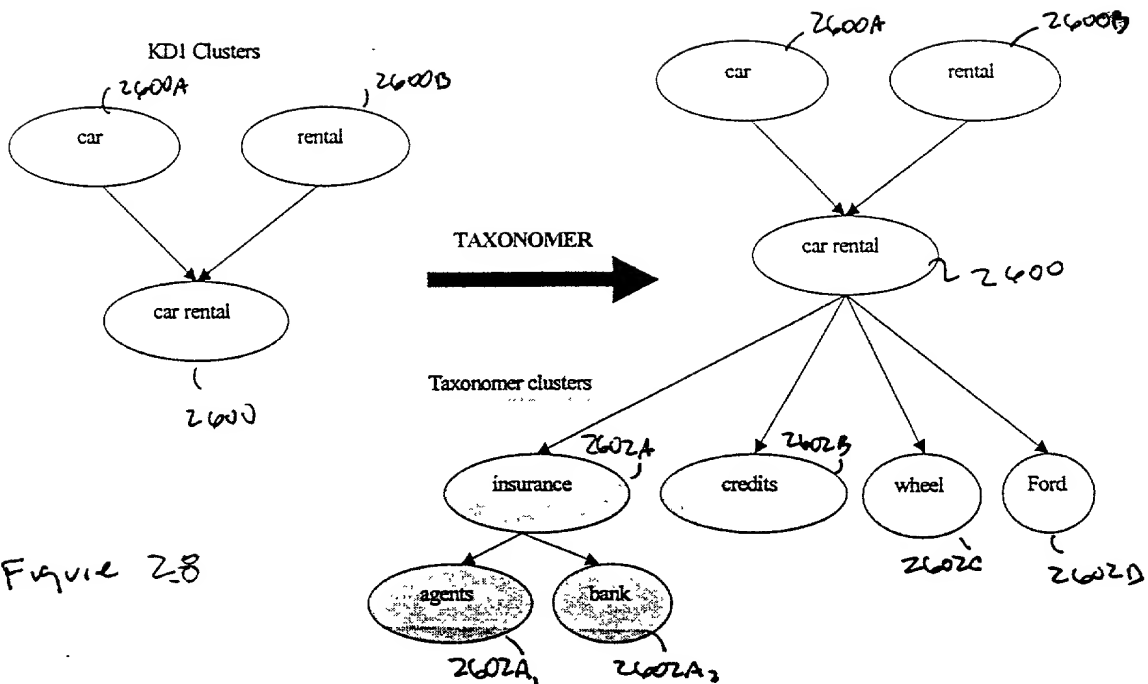


Figure 26



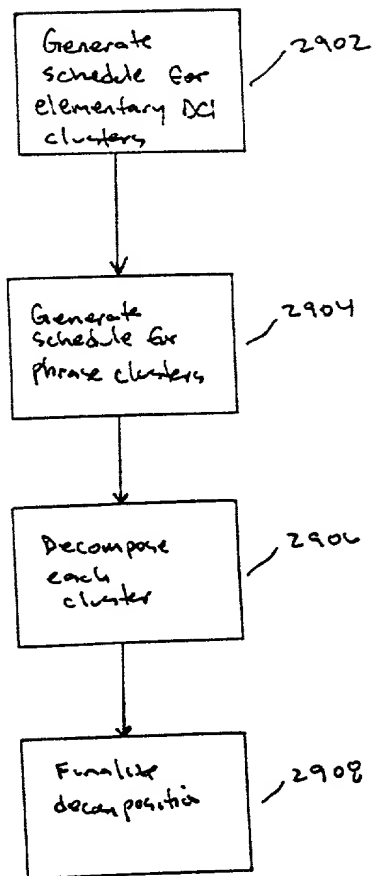


Figure 29

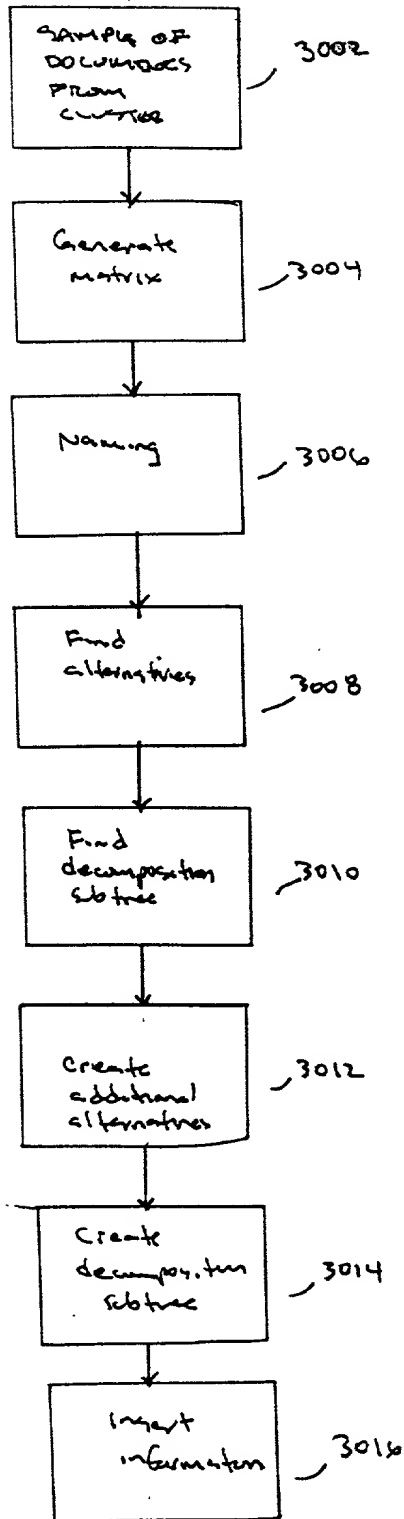


Figure 30